

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Bahija Jallal
Gregory D. Plowman
- (ii) TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
PTP04 RELATED DISORDERS
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street
Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071-2066
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0
(D) SOFTWARE: FastSEQ for Windows 2.0
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: 60/047,222
(B) FILING DATE: May 20, 1997
- (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warburg, Richard J.
(B) REGISTRATION NUMBER: 32,327
(C) REFERENCE/DOCKET NUMBER: 234/253

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (213) 489-1600
(B) TELEFAX: (213) 955-0440
(C) TELEX: 67-3510

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3580 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CCCGGGTGCC CTCCCTCAAC CTACTTATAG ACTATTTTTC TTGCTCTGCA GCATGGACCA 60
AAGAGAAATT CTGCAGAAGT TCCTGGATGA GGCCCAAAGC AAGAAAATTA CTAAAGAGGA 120
GTTTGCCAAT GAATTTCTGA AGCTGAAAAG GCAATCTACC AAGTACAAGG CAGACAAAAC 180
CTATCCTACA ACTGTGGCTG AGAAGCCCAA GAATATCAAG AAAACAGAT ATAAGGATAT 240
TTTGCCCTAT GATTATAGCC GGGTAGAAGT ATCCCTGATA ACCTCTGATG AGGATTCCAG 300
CTACATCAAT GCCAACTTCA TTAAGGGAGT TTATGGACCC AAGGCTTATA TTGCCACCCA 360
GGGTCCTTTA TCTACAACCC TCCTGGACTT CTGGAGGATG ATTTGGGAAT ATAGTGTCTT 420
TATCATTGTT ATGGCATGCA TGGAGTATGA AATGGGAAAG AAAAAGTGTG AGCGCTACTG 480
GGCTGAGCCA GGAGAGATGC AGCTGGAATT TGGCCCTTTC TCTGTATCCT GTGAAGCTGA 540
AAAAAGGAAA TCTGATTATA TAATCAGGAC TCTAAAAGTT AAGTTCAATA GTGAAACTCG 600
AACTATCTAC CAGTTTCATT ACAAGAATTG GCCAGACCAT GATGTACCTT CATCTATAGA 660
CCCTATTCTT GAGCTCATCT GGGATGTACG TTGTTACCAA GAGGATGACA GTGTTCCCAT 720
ATGCATTAC TGCAGTGCTG GCTGTGGAAG GACTGGTGTT ATTTGTGCTA TTGATTATAC 780
ATGGATGTTG CTAAAAGATG GGATAATTCC TGAGAACTTC AGTGTTTCA GTTTGATCCG 840
GGAAATGCGG ACACAGAGGC CTTCATTAGT TCAAACGCAG GAACAATATG AACTGGTCTA 900
CAATGCTGTA TTAGAACTAT TTAAGAGACA GATGGATGTT ATCAGAGATA AACATTCTGG 960
AACAGAGAGT CAAGCAAAGC ATTGTATTCC TGAGAAAAT CACACTCTCC AAGCAGACTC 1020
TTATTCTCCT AATTTACCAA AAAGTACCAC AAAAGCAGCA AAAATGATGA ACCAACAAAG 1080
GACAAAAATG GAAATCAAAG AATCTTCTTC CTTTGACTTT AGGACTTCTG AAATAAGTGC 1140
AAAAGAAGAG CTAGTTTTGC ACCCTGCTAA ATCAAGCACT TCTTTTGACT TTCTGGAGCT 1200
AAATTACAGT TTTGACAAAA ATGCTGACAC AACCATGAAA TGGCAGACAA AGGCATTTCC 1260
AATAGTTGGG GAGCCTCTTC AGAAGCATCA AAGTTTGGAT TTGGGCTCTC TTTTGTGTTG 1320
GGGATGTTCT AATTCTAAAC CTGTAAATGC AGCAGGAAGA TATTTTAATT CAAAGGTGCC 1380
AATAACACGG ACCAAATCAA CTCCTTTTGA ATTGATACAG CAGAGAGAAA CCAAGGAGGT 1440
```

```

GGACAGCAAG  GAAAACTTTT  CTTATTTGGA  ATCTCAACCA  CATGATTCTT  GTTTTGTAGA  1500
GATGCAGGCT  CAAAAAGTAA  TGCATGTTTC  TTCAGCAGAA  CTGAATTATT  CACTGCCATA  1560
TGA CTCTAAA  CACCAAATAC  GTAATGCCTC  TAATGTAAAG  CACCATGACT  CTAGTGCTCT  1620
TGGTGTATAT  TCTTACATAC  CTTTAGTGGA  AAATCCTTAT  TTTTCATCAT  GGCTCCCAAG  1680
TGGTACCAGT  TCTAAGATGT  CTCTTGATTT  ACCTGAGAAG  CAAGATGGAA  CTGTTTTTCC  1740
TTCTTCTCTG  TTGCCAACAT  CCTCTACATC  CCTCTTCTCT  TATTACAATT  CACATGATTC  1800
TTTATCACTG  AATTCTCCAA  CCAATATTTT  CTCACTATTG  AACCAGGAGT  CAGCTGTACT  1860
AGCAACTGCT  CCAAGGATAG  ATGATGAAAT  CCCCCCTCCA  CTTCTGTAC  GGACACCTGA  1920
ATCATTTTAT  GTGGTTGAGG  AAGCTGGAGA  ATTCTCACCA  AATGTTCCCA  AATCCTTATC  1980
CTCAGCTGTG  AAGGTAAAAA  TTGGAACATC  ACTGGAATGG  GGTGGAACAT  CTGAACCCAA  2040
GAAATTTGAT  GACTCTGTGA  TACTTAGACC  AAGCAAGAGT  GTAAAACTCC  GAAGTCCTAA  2100
ATCAGAACTA  CATCAAGATC  GTTCTTCTCC  CCCACCTCCT  CTCCAGAAA  GAACCTCTAG  2160
GTCTTCTTTT  CTTGCCGATG  AAGATTGTAT  GCAGGCCCAA  TCTATAGAAA  CATATTCTAC  2220
TAGCTATCCT  GACACCATGG  AAAATTCAAC  ATCTTCAAAA  CAGACACTGA  AGACTCCTGG  2280
AAAAAGTTTC  ACAAGGAGTA  AGAGTTTGAA  AATTTTGCGA  AACATGAAAA  AGAGTATCTG  2340
TAATTCTTGC  CCACCAAACA  AGCCTGCAGA  ATCTGTTTCA  TCAAATAACT  CCAGCTCATT  2400
TCTGAATTTT  GGTTTTGCAA  ACCGTTTTTC  AAAACCCAAA  GGACCAAGGA  ATCCACCACC  2460
AACTTGGAAT  ATTTAATAAA  ACTCCAGATT  TATAATAATA  TGGGCTGCAA  GTACACCTGC  2520
AAATAAACT  ACTAGAATAC  TGCTAGTTAA  AATAAGTGCT  CTATATGCAT  AATATCAAAT  2580
ATGAAGATAT  GCTAATGTGT  TAATAGCTTT  TAAAAGAAAA  GCAAATGCC  AATAAGTGCC  2640
AGTTTTCAT  TTTCATATCA  TTTGCATTGA  GTTGAAAACT  GCAAATAAAA  GTTTGTCACT  2700
TGAGCTTATG  TACAGAATGC  TATATGAGAA  ACACTTTATG  AATGGATTTA  TTTTTCATTT  2760
TTGCCAGTTA  TTTTATTTTT  CTTTACTTTT  TTTACATAAA  CATAAACTTC  AAAAGGTTTG  2820
TAAGATTTGG  ATCTCAACTA  ATTTCTACAT  TGCCAGAATA  TACTATAAAA  AGTTAAAAAA  2880
AAACTTACTT  TGTGGGTTGC  AATACAACT  GCTCTTGACA  ATGACTATTC  CCTGACAGTT  2940
ATTTTTCCT  AAATGGAGTA  TACCTTGTA  ATCTTCCCA  ATGTTGTGGA  AAAGTGAAT  3000
ATTAAGAAAA  TGAGAAATTA  TATTTATTAG  AATAAAATGT  GCAAATAATG  ACAATTATTT  3060
GAATGTAACA  AGGAATTCAA  CTGAAATCCT  GATAAGTTT  AACCAGATC  ATTAAATTAC  3120
CAATTCTAGA  AAAGTAATCA  ATGAAATATA  ATAGCTATCT  TTTGGTAGCA  AAAGATATAA  3180
ATTGTATATG  TTTATACAGG  ATCTTTCAGA  TCATGTGCAA  TTTTATCTA  ACCAATCAGA  3240
AATACTAGTT  TAAAATGAAT  TTCTATATGA  ATATGGATCT  GCCATAAGAA  AATCTAGTTC  3300
AACTCTAATT  TTATGTAGTA  AATAAATTGG  CAGGTAATTG  TTTTACAAA  GAATCCACCT  3360
GACTTCCCT  AATGCATTAA  AAATATTTTT  ATTTAAATAA  CTTTATTTAT  AACTTTTGA  3420
AACATGTAGT  ATTGTTTAAA  CATCATTTGT  TCTTCAGTAT  TTTTCATTTG  GAAGTCCAAT  3480
AGGGCAAATT  GAATGAAGTA  TTATTATCTG  TCTCTGTAG  TACAATGTAT  CCAACAGACA  3540
CTCAATAAAC  TTTTGTGTTG  TTAAAAAAA  AAAAAAAA  3580

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH:      807 amino acids
(B) TYPE:        amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY:    linear

```

(ii) MOLECULE TYPE: peptide

Met Asp Gln Arg Glu Ile Leu Gln Lys Phe Leu Asp Glu Ala Gln Ser
1 5 10 15

Lys Lys Ile Thr Lys Glu Glu Phe Ala Asn Glu Phe Leu Lys Leu Lys
20 25 30

Arg Gln Ser Thr Lys Tyr Lys Ala Asp Lys Thr Tyr Pro Thr Thr Val
35 40 45

Ala Glu Lys Pro Lys Asn Ile Lys Lys Asn Arg Tyr Lys Asp Ile Leu
50 55 60

Pro Tyr Asp Tyr Ser Arg Val Glu Leu Ser Leu Ile Thr Ser Asp Glu
65 70 75 80

Asp Ser Ser Tyr Ile Asn Ala Asn Phe Ile Lys Gly Val Tyr Gly Pro
85 90 95

Lys Ala Tyr Ile Ala Thr Gln Gly Pro Leu Ser Thr Thr Leu Leu Asp
100 105 110

Phe Trp Arg Met Ile Trp Glu Tyr Ser Val Leu Ile Ile Val Met Ala
115 120 125

Cys Met Glu Tyr Glu Met Gly Lys Lys Lys Cys Glu Arg Tyr Trp Ala
130 135 140

Glu Pro Gly Glu Met Gln Leu Glu Phe Gly Pro Phe Ser Val Ser Cys
145 150 155 160

Glu Ala Glu Lys Arg Lys Ser Asp Tyr Ile Ile Arg Thr Leu Lys Val
165 170 175

Lys Phe Asn Ser Glu Thr Arg Thr Ile Tyr Gln Phe His Tyr Lys Asn
180 185 190

Trp Pro Asp His Asp Val Pro Ser Ser Ile Asp Pro Ile Leu Glu Leu
195 200 205

Ile Trp Asp Val Arg Cys Tyr Gln Glu Asp Asp Ser Val Pro Ile Cys
210 215 220

Ile His Cys Ser Ala Gly Cys Gly Arg Thr Gly Val Ile Cys Ala Ile
225 230 235 240

Asp Tyr Thr Trp Met Leu Leu Lys Asp Gly Ile Ile Pro Glu Asn Phe

245	250	255
Ser Val Phe Ser Leu Ile Arg Glu Met Arg Thr Gln Arg Pro Ser Leu		
260	265	270
Val Gln Thr Gln Glu Gln Tyr Glu Leu Val Tyr Asn Ala Val Leu Glu		
275	280	285
Leu Phe Lys Arg Gln Met Asp Val Ile Arg Asp Lys His Ser Gly Thr		
290	295	300
Glu Ser Gln Ala Lys His Cys Ile Pro Glu Lys Asn His Thr Leu Gln		
305	310	315
Ala Asp Ser Tyr Ser Pro Asn Leu Pro Lys Ser Thr Thr Lys Ala Ala		
325	330	335
Lys Met Met Asn Gln Gln Arg Thr Lys Met Glu Ile Lys Glu Ser Ser		
340	345	350
Ser Phe Asp Phe Arg Thr Ser Glu Ile Ser Ala Lys Glu Glu Leu Val		
355	360	365
Leu His Pro Ala Lys Ser Ser Thr Ser Phe Asp Phe Leu Glu Leu Asn		
370	375	380
Tyr Ser Phe Asp Lys Asn Ala Asp Thr Thr Met Lys Trp Gln Thr Lys		
385	390	395
Ala Phe Pro Ile Val Gly Glu Pro Leu Gln Lys His Gln Ser Leu Asp		
405	410	415
Leu Gly Ser Leu Leu Phe Glu Gly Cys Ser Asn Ser Lys Pro Val Asn		
420	425	430
Ala Ala Gly Arg Tyr Phe Asn Ser Lys Val Pro Ile Thr Arg Thr Lys		
435	440	445
Ser Thr Pro Phe Glu Leu Ile Gln Gln Arg Glu Thr Lys Glu Val Asp		
450	455	460
Ser Lys Glu Asn Phe Ser Tyr Leu Glu Ser Gln Pro His Asp Ser Cys		
465	470	475
Phe Val Glu Met Gln Ala Gln Lys Val Met His Val Ser Ser Ala Glu		
485	490	495
Leu Asn Tyr Ser Leu Pro Tyr Asp Ser Lys His Gln Ile Arg Asn Ala		

500	505	510
Ser Asn Val Lys His His Asp Ser Ser Ala Leu Gly Val Tyr Ser Tyr		
515	520	525
Ile Pro Leu Val Glu Asn Pro Tyr Phe Ser Ser Trp Pro Pro Ser Gly		
530	535	540
Thr Ser Ser Lys Met Ser Leu Asp Leu Pro Glu Lys Gln Asp Gly Thr		
545	550	555
Val Phe Pro Ser Ser Leu Leu Pro Thr Ser Ser Thr Ser Leu Phe Ser		
565	570	575
Tyr Tyr Asn Ser His Asp Ser Leu Ser Leu Asn Ser Pro Thr Asn Ile		
580	585	590
Ser Ser Leu Leu Asn Gln Glu Ser Ala Val Leu Ala Thr Ala Pro Arg		
595	600	605
Ile Asp Asp Glu Ile Pro Pro Pro Leu Pro Val Arg Thr Pro Glu Ser		
610	615	620
Phe Ile Val Val Glu Glu Ala Gly Glu Phe Ser Pro Asn Val Pro Lys		
625	630	635
Ser Leu Ser Ser Ala Val Lys Val Lys Ile Gly Thr Ser Leu Glu Trp		
645	650	655
Gly Gly Thr Ser Glu Pro Lys Lys Phe Asp Asp Ser Val Ile Leu Arg		
660	665	670
Pro Ser Lys Ser Val Lys Leu Arg Ser Pro Lys Ser Glu Leu His Gln		
675	680	685
Asp Arg Ser Ser Pro Pro Pro Pro Leu Pro Glu Arg Thr Leu Glu Ser		
690	695	700
Phe Phe Leu Ala Asp Glu Asp Cys Met Gln Ala Gln Ser Ile Glu Thr		
705	710	715
Tyr Ser Thr Ser Tyr Pro Asp Thr Met Glu Asn Ser Thr Ser Ser Lys		
725	730	735
Gln Thr Leu Lys Thr Pro Gly Lys Ser Phe Thr Arg Ser Lys Ser Leu		
740	745	750

Lys Ile Leu Arg Asn Met Lys Lys Ser Ile Cys Asn Ser Cys Pro Pro
755 760 765

Asn Lys Pro Ala Glu Ser Val Gln Ser Asn Asn Ser Ser Ser Phe Leu
770 775 780

Asn Phe Gly Phe Ala Asn Arg Phe Ser Lys Pro Lys Gly Pro Arg Asn
785 790 795 800

Pro Pro Pro Thr Trp Asn Ile
805

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: The letter "Y" stands for C or T.
The letter "V" stands for A, C or G.
The letter "R" stands for A or G.
The letter "N" stands for A, C, G or T.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAYTTYTGGV RNATGRTNTG GGA

23

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: The letter "S" stands for C or G.
The letter "Y" stands for C or T.
The letter "N" stands for A, C, G
or T.
The letter "W" stands for A or T.
The letter "R" stands for A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGGCCSAYNC CNGCNSWRCA RTG

23

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in positions 4 and 6 stand
for an unspecified amino acid.
"Xaa" in position 8 stands for
either Glu or Asp.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Asp Phe Trp Xaa Met Xaa Trp Xaa
1 5

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in positions 3 and 6 stand
for an unspecified amino acid.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

His Cys Xaa Ala Gly Xaa Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CACCGTTTGA GTATTTTCTAGA TTGTGAAGAA GTCC

34

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGACTTCTTC ACAATCTGAA ATACTCGAAC GGTG

34

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	33 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CCGTTATGTG AGGAAGAGCC ACATTACAGG ACC

33

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	33 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGTCCTGTAA TGTGGCTCTT CCTCACATAA CGG

33

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	23 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGCATGCATG GAGTATGAAA TGG

23

(2) INFORMATION FOR SEQ ID NO: 12:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CGTACATCCC AGATGAGCTC AAGAATAGGG

30

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

His Cys Ser Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Ser Trp Pro Pro Ser Gly Thr Ser Ser Lys Met Ser Leu Asp Asp Leu
1 5 10 15
Pro Glu Lys Gln Asp Gly Thr Val Phe Pro Ser Ser Leu Leu Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Tyr Ser Leu Pro Tyr Asp Ser Lys His Gln Ile Arg Asn Ala Ser Asn
1 5 10 15
Val Lys His His Asp Ser Ser Ala Leu Gly Val Tyr Ser Tyr
20 25 30

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

His Thr Leu Gln Ala Asp Ser Tyr Ser Pro Asn Leu Pro Lys Ser Thr
 1 5 10 15
 Thr Lys Ala Ala Lys Met Met Asn Gln Gln Arg Thr Lys Cys
 20 25 30

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 802 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Asp Gln Arg Glu Ile Leu Gln Gln Leu Leu Lys Glu Ala Gln Lys
 1 5 10 15
 Lys Lys Leu Asn Ser Glu Glu Phe Ala Ser Glu Phe Leu Lys Leu Lys
 20 25 30
 Arg Gln Ser Thr Lys Tyr Lys Ala Asp Lys Ile Tyr Pro Thr Thr Val
 35 40 45
 Ala Gln Arg Pro Lys Asn Ile Lys Lys Asn Arg Tyr Lys Asp Ile Leu
 50 55 60
 Pro Tyr Asp His Ser Leu Val Glu Leu Ser Leu Leu Thr Ser Asp Glu
 65 70 75 80
 Asp Ser Ser Tyr Ile Asn Ala Ser Phe Ile Lys Gly Val Tyr Gly Pro
 85 90 95

Lys Ala Tyr Ile Ala Thr Gln Gly Pro Leu Ser Thr Thr Leu Leu Asp
 100 105 110
 Phe Trp Arg Met Ile Trp Glu Tyr Arg Ile Leu Val Ile Val Met Ala
 115 120 125
 Cys Met Glu Phe Glu Met Gly Lys Lys Lys Cys Glu Arg Tyr Trp Ala
 130 135 140
 Glu Pro Gly Glu Thr Gln Leu Gln Phe Gly Pro Phe Ser Ile Ser Cys
 145 150 155 160
 Glu Ala Glu Lys Lys Lys Ser Asp Tyr Lys Ile Arg Thr Leu Lys Ala
 165 170 175
 Lys Phe Asn Asn Glu Thr Arg Ile Ile Tyr Gln Phe His Tyr Lys Asn
 180 185 190
 Trp Pro Asp His Asp Val Pro Ser Ser Ile Asp Pro Ile Leu Gln Leu
 195 200 205
 Ile Trp Asp Met Arg Cys Tyr Gln Glu Asp Asp Cys Val Pro Ile Cys
 210 215 220
 Ile His Cys Ser Ala Gly Cys Gly Arg Thr Gly Val Ile Cys Ala Val
 225 230 235 240
 Asp Tyr Thr Trp Met Leu Leu Lys Asp Gly Ile Ile Pro Lys Asn Phe
 245 250 255
 Ser Val Phe Asn Leu Ile Gln Glu Met Arg Thr Gln Arg Pro Ser Leu
 260 265 270
 Val Gln Thr Gln Glu Gln Tyr Glu Leu Val Tyr Ser Ala Val Leu Glu
 275 280 285
 Leu Phe Lys Arg His Met Asp Val Ile Ser Asp Asn His Leu Gly Arg
 290 295 300
 Glu Ile Gln Ala Gln Cys Ser Ile Pro Glu Gln Ser Leu Thr Val Glu
 305 310 315 320
 Ala Asp Ser Cys Pro Leu Asp Leu Pro Lys Asn Ala Met Arg Asp Val
 325 330 335
 Lys Thr Thr Asn Gln His Ser Lys Gln Gly Ala Glu Ala Glu Ser Thr
 340 345 350

Gly Gly Ser Ser Leu Gly Leu Arg Thr Ser Thr Met Asn Ala Glu Glu
 355 360 365
 Glu Leu Val Leu His Ser Ala Lys Ser Ser Pro Ser Phe Asn Cys Leu
 370 375 380
 Glu Leu Asn Cys Gly Cys Asn Asn Lys Ala Val Ile Thr Arg Asn Gly
 385 390 395 400
 Gln Ala Arg Ala Ser Pro Val Val Gly Glu Pro Leu Gln Lys Tyr Gln
 405 410 415
 Ser Leu Asp Phe Gly Ser Met Leu Phe Gly Ser Cys Pro Ser Ala Leu
 420 425 430
 Pro Ile Asn Thr Ala Asp Arg Tyr His Asn Ser Lys Gly Pro Val Lys
 435 440 445
 Arg Thr Lys Ser Thr Pro Phe Glu Leu Ile Gln Gln Arg Lys Thr Asn
 450 455 460
 Asp Leu Ala Val Gly Asp Gly Phe Ser Cys Leu Glu Ser Gln Leu His
 465 470 475 480
 Glu His Tyr Ser Leu Arg Glu Leu Gln Val Gln Arg Val Ala His Val
 485 490 495
 Ser Ser Glu Glu Leu Asn Tyr Ser Leu Pro Gly Ala Cys Asp Ala Ser
 500 505 510
 Cys Val Pro Arg His Ser Pro Gly Ala Leu Arg Val His Leu Tyr Thr
 515 520 525
 Ser Leu Ala Glu Asp Pro Tyr Phe Ser Ser Ser Pro Pro Asn Ser Ala
 530 535 540
 Asp Ser Lys Met Ser Phe Asp Leu Pro Glu Lys Gln Asp Gly Ala Thr
 545 550 555 560
 Ser Pro Gly Ala Leu Leu Pro Ala Ser Ser Thr Thr Ser Phe Phe Tyr
 565 570 575
 Ser Asn Pro His Asp Ser Leu Val Met Asn Thr Leu Thr Ser Phe Ser
 580 585 590
 Pro Pro Leu Asn Gln Glu Thr Ala Val Glu Ala Pro Ser Arg Arg Thr
 595 600 605

Asp Asp Glu Ile Pro Pro Pro Leu Pro Glu Arg Thr Pro Glu Ser Phe
 610 615 620
 Ile Val Val Glu Glu Ala Gly Glu Pro Ser Pro Arg Val Thr Glu Ser
 625 630 635 640
 Leu Pro Leu Val Val Thr Phe Gly Ala Ser Pro Glu Cys Ser Gly Thr
 645 650 655
 Ser Glu Met Lys Ser His Asp Ser Val Gly Phe Thr Pro Ser Lys Asn
 660 665 670
 Val Lys Leu Arg Ser Pro Lys Ser Asp Arg His Gln Asp Gly Ser Pro
 675 680 685
 Pro Pro Pro Leu Pro Glu Arg Thr Leu Glu Ser Phe Phe Leu Ala Asp
 690 695 700
 Glu Asp Cys Ile Gln Ala Gln Ala Val Gln Thr Ser Ser Thr Ser Tyr
 705 710 715 720
 Pro Glu Thr Thr Glu Asn Ser Thr Ser Ser Lys Gln Thr Leu Arg Thr
 725 730 735
 Pro Gly Lys Ser Phe Thr Arg Ser Lys Ser Leu Lys Ile Phe Arg Asn
 740 745 750
 Met Lys Lys Ser Val Cys Asn Ser Ser Ser Pro Ser Lys Pro Thr Glu
 755 760 765
 Arg Val Gln Pro Lys Asn Ser Ser Ser Phe Leu Asn Phe Gly Phe Gly
 770 775 780
 Asn Arg Phe Ser Lys Pro Lys Gly Pro Arg Asn Pro Pro Ser Ala Trp
 785 790 795 800
 Asn Met